UIUCDCS-R-79-967

UILU-ENG 79 1713

An NP-Complete Matching Problem

by

David A. Plaisted

April 1979

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN · URBANA, ILLINOIS

Corrections to University of Illinois Departmental Report "An NP-Complete Matching Problem" by David A. Plaisted

page 1 line 13 "vertex of E1" should be "edge of E1."

page 9 line -2 "q edges" should be "q triples"

An NP-Complete Matching Problem

by

David A. Plaisted
222 Digital Computer Lab
Department of Computer Science
University of Illinois
Urbana, Illinois 61801

This research was partially supported by the National Science Foundation under grant MCS 77-22830.

Abstract

We show that a restricted form of the perfect matching problem for bipartite graphs is NP-complete. The restriction involves partitions of the vertices of the graph. This problem is still NP-complete if the degrees of the vertices are restricted to be 3 or less. For degrees restricted to 2 or less, a polynomial time algorithm exists.

1. Introduction

The general perfect matching problem can be solved in polynomial time [4], but certain related problems are NP-complete [2, 3]. We show that a more restricted version of the perfect matching problem is still NP-complete. Much of this work was done together with Shmuel Zaks.

Suppose G = (V, E) is an undirected graph, where V is the set of vertices and E is the set of edges. We represent an edge joining v_1 and v_2 by the set $\{v_1, v_2\}$ for vertices v_1 and v_2. Suppose G is a bipartite graph. That is, $V = V_1 \cup V_2$, $V_1 \cap V_2 = \emptyset$, and every edge joins a vertex in V_1 with a vertex in V_2. We represent such a bipartite graph by $G = (V_1, V_2, E)$.

The perfect matching problem is to determine if there is a subset E1 of E such that every vertex of G is incident with exactly one vertex of E1. We are interested in the following problem, which we call "perfect matching with node partitions."

> Given an undirected bipartite graph $G = (V_1, V_2, E)$ and given partitions P_1 and P_2 of V_1 and V_2, respectively, to determine if there is a subset E1 of E such that E1 is a perfect matching for G and such that no distinct edges $\{v_1, v_2\}$ and $\{w_1, w_2\}$ of E1 have the property that v_1 and w_1 are in the same block of P_1 and v_2 and w_2 are in the same block of P_2.

We say that such a matching E1 is a perfect matching <u>consistent</u> <u>with</u> the node partitions P_1 and P_2. This problem is a specialization of the 3-dimensional matching problem [3] and of the restricted matching problem treated in [2]. Both of these latter problems are known to be NP-complete.

A motivation for the perfect matching problem with node partitions is the following: Let V_1 be a set of people and let V_2 be a set of offices in various committees. Let E be a set of pairs $\{v_1, v_2\}$ indicating that person v_1 is eligible for office v_2. We may partition V_1 into families and V_2 into committees. The problem is to assign people to committees so that no two people from the same family are on the same committee. It is easy to see that this is an instance of the perfect matching problem with node partitions.

2. NP-Completeness

We now show that this restricted perfect matching problem is NP-complete. Clearly it is in NP. We show that it is NP-hard by reducing from the one-in-three 3-SAT problem, known to be NP-complete [5]. The final form of this reduction is mostly due to S. Zaks.

The one-in-three 3SAT problem is as follows: Given a set X of variables and a set S of clauses over X such that for all $C \in S$, C has 3 literals, to determine if there is a 1-truth assignment (a truth assignment making exactly one literal in each clause true).

First we reduce the one-in-three 3SAT problem to a related version of the one-in-three 3SAT problem. The difference is that for all variables x in X, at most 3 clauses in S contain either x or $\bar{x}$. However, some clauses in S may have only 2 literals. This reduction proceeds as follows: Suppose x or $\bar{x}$ appears in $C_1, C_2, \ldots, C_m$, for clauses $C_i \in S$, and x or $\bar{x}$ appears in no other clauses of S. Suppose $m > 3$. Introduce new variables $y_1, y_2, \ldots, y_m$. Let C_i' be C_i with x replaced by y_i. Let S' be with C_i replaced by C_i', for $1 \leq i \leq m$, and with the additional clauses

$\bar{y}_1 \vee y_2$, $\bar{y}_2 \vee y_3$, $\bar{y}_3 \vee y_4$, $\ldots$, $\bar{y}_{m-1} \vee y_m$, $\bar{y}_m \vee y_1$ added. Now it is clear that S' has a 1-truth assignment iff S does, since any interpretation satisfying the clauses $\bar{y}_1 \vee y_2$, $\bar{y}_2 \vee y_3$, $\ldots$, $\bar{y}_m \vee y_1$ must give all y_i the same truth value. Also, each y_i appears in exactly three clauses. Hence by repeating this step, we obtain a set T of clauses which has a 1-truth assignment iff S does, and such that for each variable x, at most 3 clauses contain x or $\bar{x}$.

We now reduce the restricted one-in-three 3SAT problem to the restricted matching problem. Suppose S is a set of clauses over the variables X. Suppose each clause C in S has 3 or fewer literals, and for each x in X, at most 3 clauses of S contain either x or $\bar{x}$. We construct a bipartite graph $G = (V_1, V_2, E)$ and partitions P_1 of V_1 and P_2 of V_2 such that G has a matching as desired iff S has a 1-truth assignment. This construction of G from S can be done in polynomial time, as will be clear. In fact, the construction can be done in linear time.

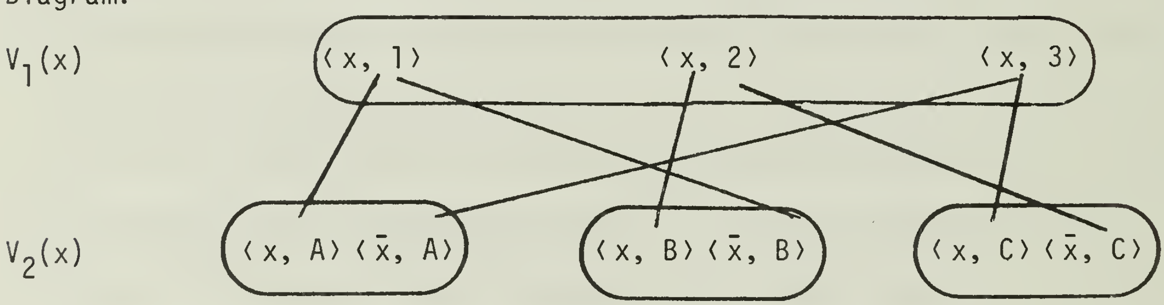
Let L be $X \cup \{\bar{x}\colon x \in X\}$. Thus L is the set of <u>literals</u> of S. We say $L \in C$ if L is a literal appearing in clause C. Thus $\bar{x} \in \{\bar{x} \vee y \vee z\}$ but $x \notin \{\bar{x} \vee y \vee z\}$. For literals L in L we define $\bar{L}$ in the following way: If L is x then $\bar{L}$ is $\bar{x}$; if L is $\bar{x}$ then $\bar{L}$ is x. We call $\bar{L}$ the <u>complement</u> of L. Let S_x be $\{C \in S\colon x \in C \text{ or } \bar{x} \in C\}$ for $x \in X$.

The graph G is the "union" of graphs obtained from the clauses and variables in S. We define the union of graphs $G' = (V_1', V_2', E')$ and $G'' = (V_1'', V_2'', E'')$ to be the graph $G' \cup G'' = (V_1' \cup V_1'', V_2' \cup V_2'', E' \cup E'')$. For each variable x in X, we define a graph G(x) in a manner to be described, and for each clause C in S we define a graph G(C) in a manner to be described. The final graph G is then $(\bigcup_{x \in X} G(x)) \cup (\bigcup_{C \in S} G(C))$. Also, the

partitions P_1 and P_2 are defined by $P_1 = (\bigcup_{x \in X} P_1(x)) \cup (\bigcup_{C \in S} P_1(C))$ and $P_2 = (\bigcup_{x \in X} P_2(x)) \cup (\bigcup_{C \in S} P_2(C))$ where $P_1(x)$, $P_1(C)$, $P_2(x)$, and $P_2(C)$ remain to be defined.

We define $G(C)$ and $G(x)$ by cases. Suppose $x \in X$ and $|S_x| = 3$. Let $\{A, B, C\}$ be the set of three (distinct) clauses in which x or $\bar{x}$ appears. Then $G(x) = (V_1(x), V_2(x), E(x))$ where $V_1(x) = \{\langle x, 1\rangle, \langle x, 2\rangle, \langle x, 3\rangle\}$, $V_2(x) = \{\langle x, A\rangle, \langle \bar{x}, A\rangle, \langle x, B\rangle, \langle \bar{x}, B\rangle, \langle x, C\rangle, \langle \bar{x}, C\rangle\}$, and $E(x) = \{\{\langle x, 1\rangle, \langle x, A\rangle\}, \{\langle x, 1\rangle, \langle \bar{x}, B\rangle\}, \{\langle x, 2\rangle, \langle x, B\rangle\}, \{\langle x, 2\rangle, \langle \bar{x}, C\rangle\}, \{\langle x, 3\rangle, \langle x, C\rangle\}, \{\langle x, 3\rangle, \langle \bar{x}, A\rangle\}\}$. Note that $G(x)$ may differ depending on which order we take A, B, and C. Any order will do. Also, $P_1(x) = \{V_1(x)\}$ and $P_2(x) = \{\{\langle x, A\rangle, \langle \bar{x}, A\rangle\}, \{\langle x, B\rangle, \langle \bar{x}, B\rangle\}, \{\langle x, C\rangle, \langle \bar{x}, C\rangle\}\}$.

Diagram:



$G(x)$

The ovals illustrate the blocks of the partition. Note that there are only two restricted matchings of this graph. One of them leaves the vertices $\langle x, A\rangle$, $\langle x, B\rangle$, and $\langle x, C\rangle$ free, and the other one leaves $\langle \bar{x}, A\rangle$, $\langle \bar{x}, B\rangle$, and $\langle \bar{x}, C\rangle$ free. (For example, if $\langle x, 1\rangle$ matches $\langle \bar{x}, B\rangle$ then $\langle x, 2\rangle$ can't match $\langle x, B\rangle$ so $\langle x, 2\rangle$ matches $\langle \bar{x}, C\rangle$ and so on.) The former matching corresponds to an interpretation making x true and the latter matching corresponds to an interpretation making x false.

If $|S_x| = 2$ or $|S_x| = 1$, similar constructions for $G(x)$ are used.

Suppose $|S_x| = 2$ and $S_x = \{A, B\}$. We then have $V_1(x) = \{\langle x, 1\rangle, \langle x, 2\rangle\}$ and $V_2(x) = \{\langle x, A\rangle, \langle \bar{x}, A\rangle, \langle x, B\rangle, \langle \bar{x}, B\rangle\}$ and $\mathcal{P}_1(x) = \{V_1(x)\}$ and $\mathcal{P}_2(x) = \{\{\langle x, A\rangle, \langle \bar{x}, A\rangle\}, \{\langle x, B\rangle, \langle \bar{x}, B\rangle\}\}$. Also, $E(x) = \{\{\langle x, 1\rangle, \langle x, A\rangle\}, \{\langle x, 1\rangle, \langle \bar{x}, B\rangle\}, \{\langle x, 2\rangle, \langle x, B\rangle\}, \{\langle x, 2\rangle, \langle \bar{x}, A\rangle\}\}$.

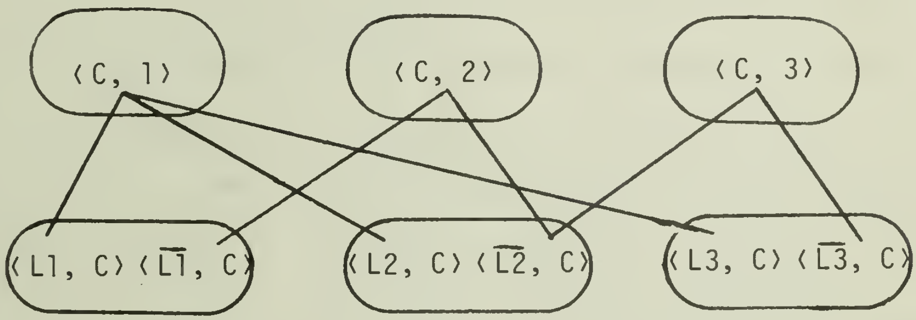
Suppose $|S_x| = 1$ and $S_x = \{A\}$. We then have $V_1(x) = \{\langle x, 1\rangle\}$ and $V_2(x) = \{\langle x, A\rangle, \langle \bar{x}, A\rangle\}$ and $\mathcal{P}_1(x) = \{V_1(x)\}$ and $\mathcal{P}_2(x) = \{V_2(x)\}$. Also, $E(x) = \{\{\langle x, 1\rangle, \langle x, A\rangle\}, \{\langle x, 1\rangle, \langle \bar{x}, A\rangle\}\}$.

We define $G(C)$ for $C \in S$ as follows: Suppose $C = L1 \vee L2 \vee L3$ for $L1, L2, L3 \in \mathcal{L}$. Then $V_1(C) = \{\langle C, 1\rangle, \langle C, 2\rangle, \langle C, 3\rangle\}$ and $V_2(C) = \{\langle L1, C\rangle, \langle \overline{L1}, C\rangle, \langle L2, C\rangle, \langle \overline{L2}, C\rangle, \langle L3, C\rangle, \langle \overline{L3}, C\rangle\}$. Also, $\mathcal{P}_1(C) = \{\{\langle C, 1\rangle\}, \{\langle C, 2\rangle\}, \{\langle C, 3\rangle\}\}$ and $\mathcal{P}_2(C) = \{\{\langle L1, C\rangle, \langle \overline{L1}, C\rangle\}, \{\langle L2, C\rangle, \langle \overline{L2}, C\rangle\}, \{\langle L3, C\rangle, \langle \overline{L3}, C\rangle\}\}$. Finally, $E(C) = \{\{\langle C, 1\rangle, \langle L1, C\rangle\}, \{\langle C, 1\rangle, \langle L2, C\rangle\}, \{\langle C, 1\rangle, \langle L3, C\rangle\}, \{\langle C, 2\rangle, \langle \overline{L1}, C\rangle\}, \{\langle C, 2\rangle, \langle \overline{L2}, C\rangle\}, \{\langle C, 3\rangle, \langle \overline{L2}, C\rangle\}, \{\langle C, 3\rangle, \langle \overline{L3}, C\rangle\}\}$.

Diagram:

$V_1(C)$

$V_2(C)$



$G(C)$

The vertex $\langle C, 1\rangle$ can match any one of the vertices $\langle L1, C\rangle$, $\langle L2, C\rangle$, $\langle L3, C\rangle$. The vertices $\langle C, 2\rangle$ and $\langle C, 3\rangle$ can then match $\langle \bar{L}, C\rangle$ for all $L \in C$ such that $\langle C, 1\rangle$ does not match $\langle L, C\rangle$. We cannot have $\langle C, 1\rangle$

matching ⟨ L1, C⟩ and ⟨ C, 2⟩ matching ⟨ $\overline{\text{L1}}$, C⟩ because only one of ⟨ L1, C⟩ and ⟨ $\overline{\text{L1}}$, C⟩ will be left free, by allowable matchings on G(x) for x = L or $\bar{x}$ = L. If ⟨ C, 1⟩ matches ⟨ L1, C⟩, this corresponds to L1 being true in an interpretation and L2 and L3 being false (as required by one-in-three 3-satisfiability), and similarly if ⟨ C, 1⟩ matches ⟨ L2, C⟩ or ⟨ L3, C⟩. The construction for C = L1 $\vee$ L2 or for C = L1 is similar.

Finally, $G = (\bigcup_{x \in X} G(x)) \cup (\bigcup_{C \in S} G(C))$ as stated earlier. We now show that G has a matching consistent with the constraints represented by P_1 and P_2 iff S has a 1-truth assignment. Assume G has such a matching. Then each G(x) has a matching consisting of $|S_x|$ edges. Also, this matching is consistent with the node partitions $P_1(x)$ and $P_2(x)$ of G(x). Each such matching corresponds to a truth value for x as indicated above. Consider the interpretation I of X giving all $x \in X$ truth values as specified by these matchings of G(x), for $x \in X$. Since each G(C) has a matching, at least one literal of C must be true in I. By the construction of G(C), if such a matching exists then the other two literals of C must be false in I. Hence I is a 1-truth assignment for S. Conversely, if a 1-truth assignment I for S exists, a matching for G(x) exists corresponding to the truth value of x in I, for $x \in X$. This matching will be consistent with the node partitions $P_1(x)$ and $P_2(x)$ of G(x). Also, matchings for G(C) can be obtained since exactly one literal of C is true in I. These matchings can be combined to obtain a matching for G consistent with the node partitions. This completes the proof.

3. Further restrictions

Note that the above graph G has the property that no node has

degree larger than three and that no block of P_1 or P_2 has more than three elements. We now indicate how this result can still be obtained if we restrict the blocks of P_1 and P_2 to have no more than two elements, while maintaining the bound of three on node degrees. The idea is to define G as above except that G(x) is defined differently for $x \in X$. In particular, if $|S_x| = 3$, let A, B, and C be the three (distinct) clauses in which x or $\bar{x}$ appears. Then $G(x) = (V_1(x), V_2(x), E(x))$ where

$V_1(x) = \{x, \bar{x}\} \underline{x} \{1, 2, 3\}$ and

$V_2(x) = (\{x, \bar{x}\} \underline{x} \{A, B, C\}) \cup \{x\} \underline{x} \{4, 5, 6\}$.

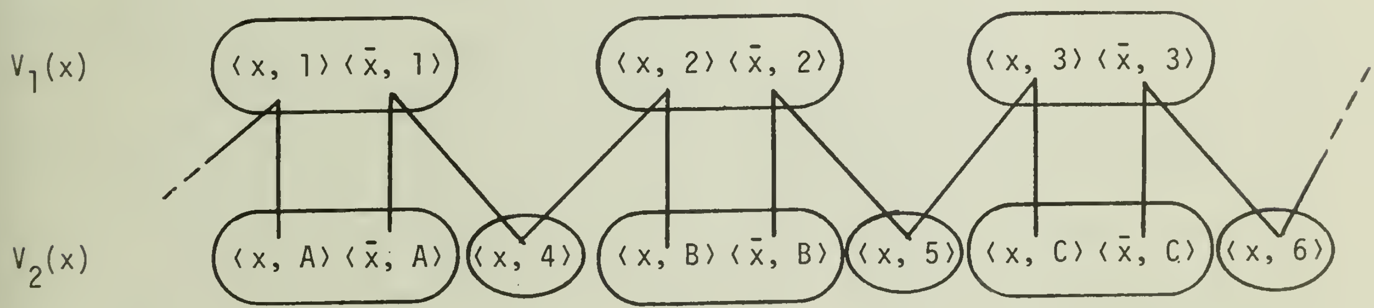
Here $\underline{x}$ indicates Cartesian product. Also,

$$\begin{aligned} E(x) = \{&\{\langle x, 1\rangle, \langle x, A\rangle\}, \{\langle \bar{x}, 1\rangle, \langle \bar{x}, A\rangle\}, \{\langle x, 2\rangle, \langle x, B\rangle\}, \\ &\{\langle \bar{x}, 2\rangle, \langle \bar{x}, B\rangle\}, \{\langle x, 3\rangle, \langle x, C\rangle\}, \{\langle \bar{x}, 3\rangle, \langle \bar{x}, C\rangle\}\} \\ \cup \{&\{\langle \bar{x}, 1\rangle, \langle x, 4\rangle\}, \{\langle x, 2\rangle, \langle x, 4\rangle\}, \{\langle \bar{x}, 2\rangle, \langle x, 5\rangle\}, \\ &\{\langle x, 3\rangle, \langle x, 5\rangle\}, \{\langle \bar{x}, 3\rangle, \langle x, 6\rangle\}, \{\langle x, 1\rangle, \langle x, 6\rangle\}\} \end{aligned}$$

and $P_1(x) = \{\{\langle x, 1\rangle, \langle \bar{x}, 1\rangle\}, \{\langle x, 2\rangle, \langle \bar{x}, 2\rangle\}, \{\langle x, 3\rangle, \langle \bar{x}, 3\rangle\}\}$

and $P_2(x) = \{\{\langle x, A\rangle, \langle \bar{x}, A\rangle\}, \{\langle x, B\rangle, \langle \bar{x}, B\rangle\}, \{\langle x, C\rangle, \langle \bar{x}, C\rangle\},$
$\{\langle x, 4\rangle\}, \{\langle x, 5\rangle\}, \{\langle x, 6\rangle\}\}$.

Diagram:



G(x)

The idea is that if $\langle x, 1\rangle$ matches $\langle x, A\rangle$ then $\langle \bar{x}, 1\rangle$ must match $\langle x, 4\rangle$. Therefore $\langle x, 2\rangle$ must match $\langle x, B\rangle$ and so on. If $\langle x, 1\rangle$ matches $\langle x, 6\rangle$ then $\langle \bar{x}, 3\rangle$ must match $\langle \bar{x}, C\rangle$ so $\langle x, 3\rangle$ must match $\langle x, 5\rangle$ and so on. In the former case, the nodes $\langle \bar{x}, A\rangle$, $\langle \bar{x}, B\rangle$, and $\langle \bar{x}, C\rangle$ of $V_2(x)$ are free, and in the latter case $\langle x, A\rangle$, $\langle x, B\rangle$, and $\langle x, C\rangle$ are free. Therefore the same argument as before holds; the former case corresponds to an interpretation making x false and the latter case corresponds to an interpretation making x true. However, with G(x) as defined here, no block of P_1 or P_2 has more than 2 vertices. If $|S_x| = 2$ or $|S_x| = 1$ a similar construction is used.

Reducing the blocks of P_1 and P_2 to contain only one vertex yields the usual perfect matching problem for which a polynomial time algorithm is known [4] regardless of the degrees of the vertices. We now show that reducing the degrees of the vertices to 2 permits a polynomial time solution regardless of the sizes of the blocks of P_1 and P_2.

Let $G = (V_1, V_2, E)$ be a bipartite graph in which the degree of each vertex is 2 or less. Let P_1 and P_2 be arbitrary partitions of V_1 and V_2, respectively. We show how to construct in polynomial time a set S of 2-literal clauses such that S is satisfiable iff G has a matching consistent with the node partitions. Since 2-satisfiability has a polynomial time solution [1], the given problem also has a polynomial time solution if the degrees are restricted to be 2 or less.

The set S of clauses has a propositional variable e for each edge e of G. Also, if distinct edges e_1 and e_2 meet a common vertex, the clauses $e_1 \vee e_2$ and $\bar{e}_1 \vee \bar{e}_2$ are in S. This signifies that exactly

one of e_1 and e_2 must be included in a perfect matching. If distinct edges e_1 and e_2 join vertices in the same block of P_1 with vertices in the same block of P_2, then $\bar{e}_1 \vee \bar{e}_2$ is in S. This signifies that at most one of e_1 and e_2 can be in a matching consistent with the node partitions. No other clauses are in S. It should be clear that S is satisfiable iff G has a perfect matching consistent with the node partitions P_1 and P_2.

4. Comparison with other matching problems

The general problem of perfect matching with node partitions is a restriction of multiple choice matching [2] because the sets of edges in multiple choice matching are arbitrary, whereas for the former problem the sets of edges must correspond to node partitions. (Multiple choice matching specifies subsets of the edges, each of which can contain at most one edge in the matching). In fact, for node partitions with at most two nodes per block of the partition, our construction yields edge sets containing at most two edges. (Two edges are in the same edge set if they join nodes in the same block of P_1 with nodes in the same block of P_2.)

We now show that perfect matching with node partitions is a restriction of 3-dimensional matching. The 3-dimensional matching problem is the following: Given a set $M \subseteq W \times Y \times Z$ where W, Y, and Z are disjoint sets having the same number q of elements, does M contain a matching containing q edges? A matching is a subset M' of M such that no two elements of M' agree in any co-ordinate.

Given a graph $G = (V_1, V_2, E)$, with $|V_1| = |V_2|$ and given partitions P_1 and P_2 of V_1 and V_2, respectively, we construct a problem closely related to 3-dimensional matching. Let W be V_1, let Y be V_2, and let Z be $P_1 \underline{x} P_2$. Let M be $\{\langle v_1, v_2, \langle p_1, p_2\rangle\rangle : v_1 \in V_1, v_2 \in V_2, \{v_1, v_2\} \in E$, and $p_1 \in P_1, p_2 \in P_2$ and $v_1 \in p_1, v_2 \in p_2\}$. That is, p_1 and p_2 are the blocks of P_1 and P_2 containing v_1 and v_2, respectively. Now, a matching M' of M corresponds to a matching of G in which no two distinct edges join vertices in the same block of P_1 with vertices in the same block of P_2. Hence a matching M' exists with $|M'| = |V_1| = |V_2|$ iff G has a perfect matching consistent with the node partitions. However, we do not have $|M'| = |Z|$ since it may be that $|V_1| \neq |Z|$ and $|V_2| \neq |Z|$. Despite this, it is easy to use this modified form of 3-dimensional matching to show that the general 3-dimensional matching problem is NP-complete.

References

[1] Cook, S. A. The complexity of theorem proving procedures, Proc. 3rd Annual ACM Symposium on Theory of Computing (1971) 151-158.

[2] Itai, A. Rodeh, M. and Tanimoto, S. L. Some matching problems for bipartite graphs, J. ACM 25(1978) 517-525.

[3] Karp, R. M. Reducibility among combinatorial problems, in Miller, R. E. and Thatcher, J. W., Complexity of Computer Computations (Plenum Press, New York, 1972) 85-104.

[4] Lawler, E. Combinatorial Optimization (Holt, Rinehart, and Winston, New York, 1976).

[5] Schaefer, T. J. The complexity of satisfiability problems, Proc. 10th Annual ACM Symposium on Theory of Computing (1978) 216-226.

BIBLIOGRAPHIC DATA SHEET	1. Report No. UIUCDCS-R-79-967	2.	3. Recipient's Accession No.
4. Title and Subtitle An NP-Complete Matching Problem			5. Report Date April 1979
			6.
7. Author(s) David A. Plaisted			8. Performing Organization Rept. No.
9. Performing Organization Name and Address Dept. of Computer Science 222 Digital Computer Lab University of Illinois Urbana, Illinois 61801			10. Project/Task/Work Unit No.
			11. Contract/Grant No. NSF MCS 77-22830
12. Sponsoring Organization Name and Address National Science Foundation Washington, D.C.			13. Type of Report & Period Covered
			14.

15. Supplementary Notes

16. Abstracts

We show that a restricted form of the perfect matching problem for bipartite graphs is NP-complete. The restriction involves partitions of the vertices of the graph. This problem is still NP-complete if the degrees of the vertices are restricted to be 3 or less. For degrees restricted to 2 or less, a polynomial time algorithm exists.

17. Key Words and Document Analysis. 17a. Descriptors

perfect matching, 3-dimensional matching, graph algorithms, NP-complete.

17b. Identifiers/Open-Ended Terms

17c. COSATI Field/Group

18. Availability Statement	19. Security Class (This Report) UNCLASSIFIED	21. No. of Pages
	20. Security Class (This Page) UNCLASSIFIED	22. Price

FORM NTIS-35 (10-70) USCOMM-DC 40329-P71